NMC Ultrascale Systems Research Center

New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY EST. 1943

Research Park Building
Suite 200, Camelina Conference Room #26
Wednesday, November 28, 2012, 2:30 - 3:30 pm

Xin Yuan, Florida State University

Title 1: A New Design of RDMA-based Small Message Channels for InfiniBand Clusters

Title 2: LFTI: A Performance Metric for Evaluating Interconnect Topology and Routing Designs for Large-Scale HPC Systems

This talk has two independent parts. The first part is on a new design of RDMA-based small message channels for InfiniBand clusters that improves over the MVAPICH RDMA-based small message channels. I will first describe a technique that eliminates persistent buffer association, a scheme used in MVAPICH that not only results in significant memory requirement, but also imposes restrictions in memory management. Building upon this technique, we develop a novel shared RDMA-based small message channel design that allows MPI processes on the same SMP node to share small message channels, which greatly reduces the number of small message channels needed for an MPI program on clusters with SMP nodes. Our techniques considerably improve the scalability and reduce memory requirement in comparison to MVAPICH, allowing RDMA-based small message channels to be used by a much larger number of MPI processes. The experimental results demonstrate that our techniques achieve the improvements without adding noticeable overheads or sacrificing the performance benefits of RDMA.

In the second part of the talk, I will first discuss the limitations of the current approaches for evaluating topology and routing designs for HPC systems. I will then present a new performance metric, called LANL-FSU Throughput Indices (LFTI), that we develop for evaluating interconnect topology and routing designs. Unlike other existing performance metrics that are not very indicative of the application level performance, LFTI directly reflects the topology/routing's support for common communication patterns. LFTI can be derived from topology and routing specification without detailed simulation,which allows for rapid and comprehensive evaluation and comparison of different topology and routing designs. Finally, I will present preliminary results of our comparison of some competing topologies for the next generation supercomputers including fat-tree, dragonfly, bristled hypercube, 6D torus, 3D torus, and jellyfish.

Xin Yuan is currently a full Professor in the Department of Computer Science at Florida State University. His research interests include parallel and distributed systems, interconnection networks, communication optimizations, and networking. He obtained his B.S. and M.S degrees in Computer Science from Shanghai Jiaotong University in 1989 and 1992, respectively. He earned his Ph.D degree in Computer Science from the University of Pittsburgh in 1998. The STAR-MPI software package that he and his students developed has been incorporated in the MPI stack of the IBM Blue Gene/P system. Dr. Yuan is currently serving on the Editorial Boards of several international journals. He has also served as the Program Chairs and vice-Chairs for several international conferences and workshops such as the International Conference on Parallel Processing (ICPP) and the IEEE International Conference on High Performance Computing (HiPC), and as Program Committee Members for many international conferences and workshops. He is a senior member of ACM and IEEE.

http://newmexicoconsortium.org/usrc

usrc@newmexicoconsortium.org

NM STATE UNIVERSITY

NEW MEXICO TECH
SCIENCE · ENGINEERING · RESEARCH · UNIVERSITY

The University of New Mexico

The USRC is a collaboration between the New Mexico Consortium (NMC) and LANL to engage universities and industry nationally in support of exascale research. DOE through the Office of Science ASCR program and the NNSA ASC program is working towards a national Exascale Computing program. This program has as a target to build an Exascale Computing Platform and Environment between 2018 and 2020.